Modeling the segment polarity gene network

First: System is biologically defined; known expression patterns

Input: segment polarity genes

Hypotheses:

continuous model: transcription factors act as enzymes

Boolean model: mRNA and protein activity is switch-like

Validation: reproduces known gene expression patterns.

Explored: changes in kinetic parameters

mutations

changes in initial conditions

Insight: topology is a main source of robustness.

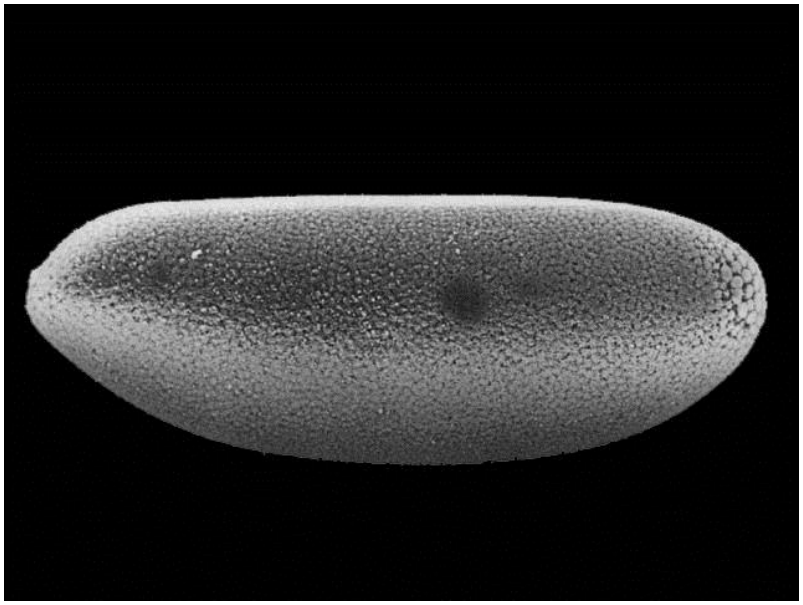
G. von Dassow et al., Nature 406, 188 (2000)

R. Albert, H. G. Othmer, Journ. Theor. Biol. 223, 1 (2003)

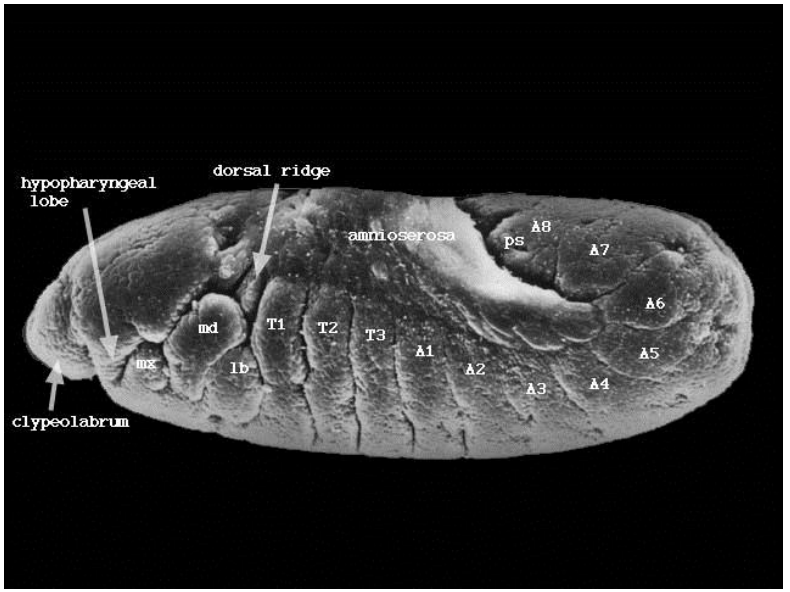
M. Chaves, R. Albert, E. Sontag Journ. Theor. Bio. 235, 431 (2005).

M. Chaves, E. Sontag, R. Albert, IEE Proc. Systems Biology 153, 154 (2006).

Segmentation of the fruit fly embryo



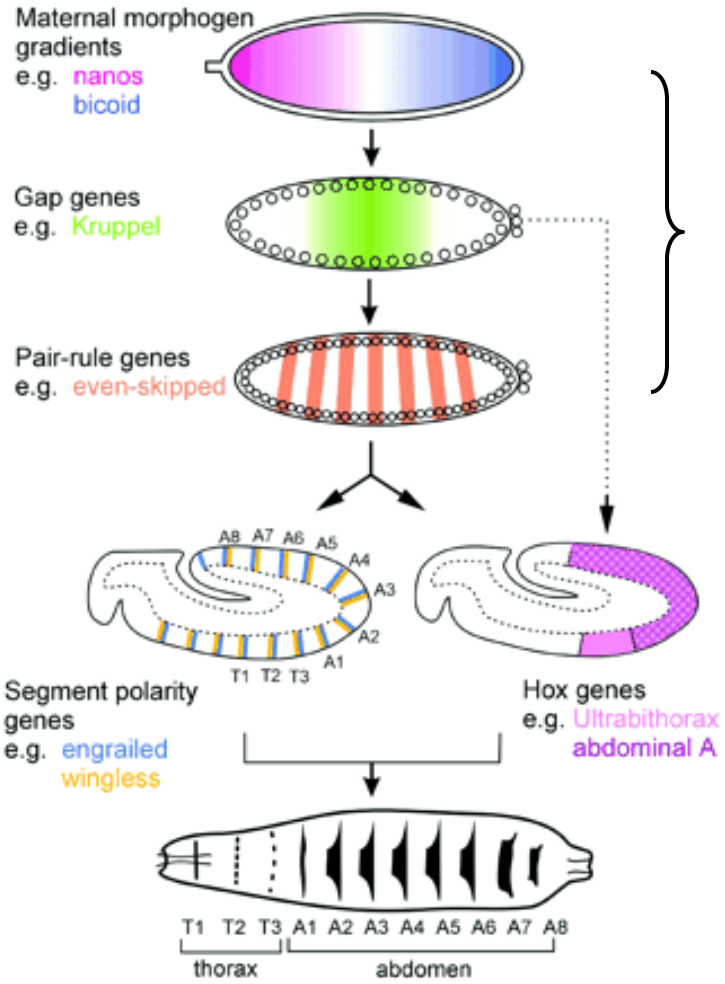
Syncytial blastoderm, 1h



End of gastrulation, 7h

- Cell differentiation is based on differential gene expression.
- The segment polarity genes determine and maintain the parasegment borders.

Segmentation is governed by a cascade of genes



Transient gene products, initiate the next step then disappear.

Ex: construct a hypothetical scenario for generating the periodic pattern

The role of the segment polarity genes

- The segment polarity genes are initiated by the pair-rule genes
- Several segment polarity genes are expressed (active) in stripes that are repeated in every fourth cell.
- These genes interact via a complex regulatory network.
- The expression pattern of the segment polarity genes is maintained for 3 hours.
- The parasegment borders appear between the cells expressing the two most important segment polarity genes, *engrailed* and *wingless*.

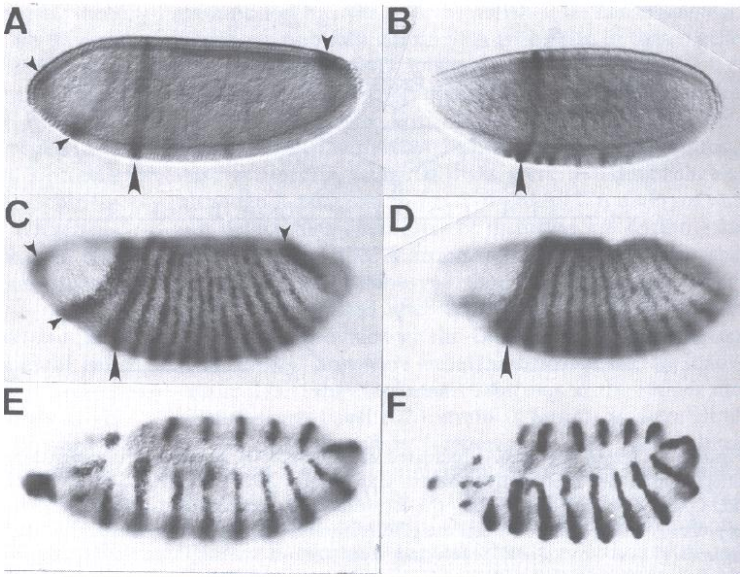
Segment polarity genes

Genes		Proteins
•*wingless (wg)*	→	Wingless protein (WG) - secreted
•*hedgehog (hh)*	→	Hedgehog protein (HH) - secreted
•*engrailed (en)*	→	Engrailed protein (EN) - transcription factor
•*patched (ptc)*	→	Patched protein (PTC) - receptor
•*cubitus interruptus (cid)*	→	Cubitus activator (CID) - transcription factor
	→	Cubitus repressor (CN) - transcription factor

Gene products form a network that maintains a gene expression pattern initiated in an earlier stage.

Evolution of gene expression patterns

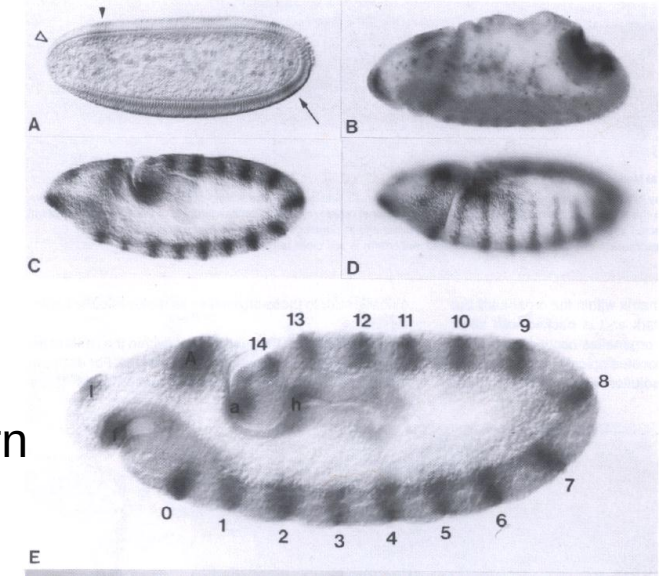
en *hh* *wg*

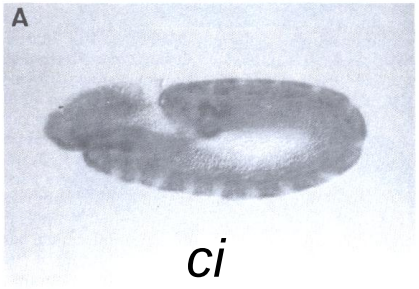


early stages
2:50 h

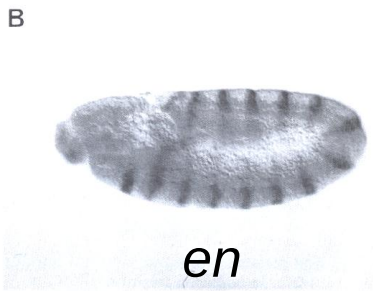
pre-pattern
3:00-3:30 h

stable pattern
4:20-7:20 h





ci

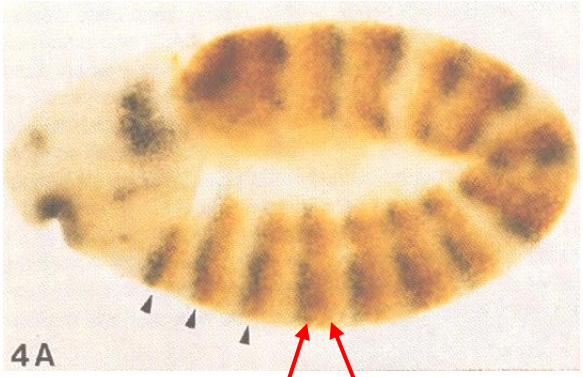


en

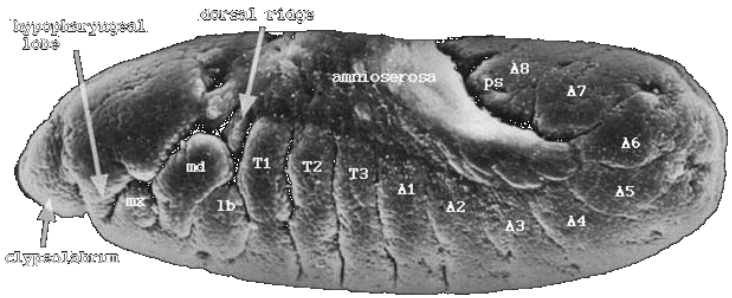
ptc

3:30 h

Wild type, stable gene patterns



wg en

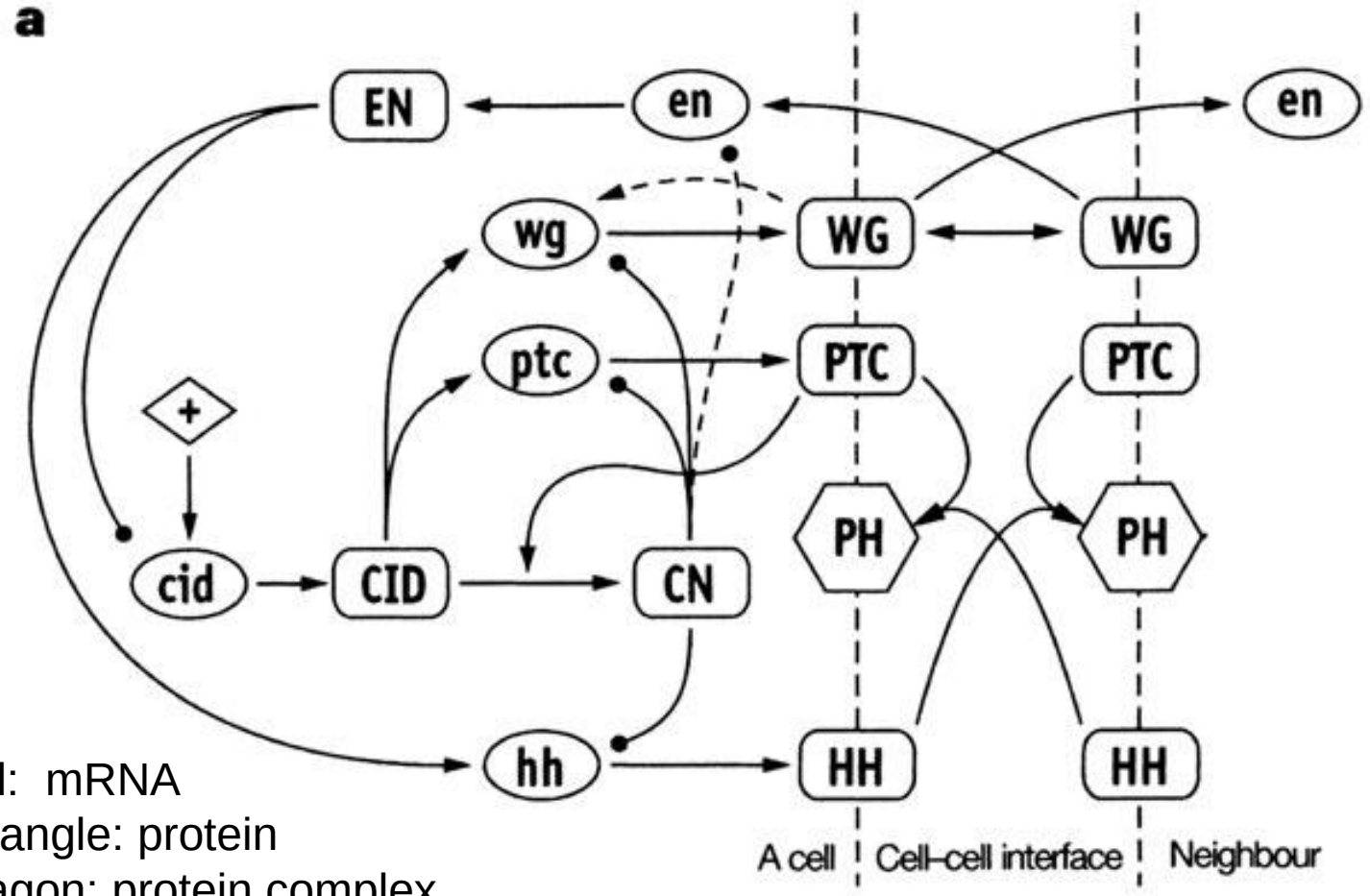


wg en

- *en* is expressed in the anterior part of the parasegment.
- *wg* is expressed in the posterior part of the parasegment.
- parasegmental grooves form between the *wg* and *en* stripes.
- *two ptc* stripes in each parasegment.
- *ci* pattern is complementary to that of *en.*

Ex. Draw a one-dimensional representation of the initial and final gene expression patterns.

Gene interaction network in the continuous model



oval: mRNA
rectangle: protein
octagon: protein complex
inhibition
transport

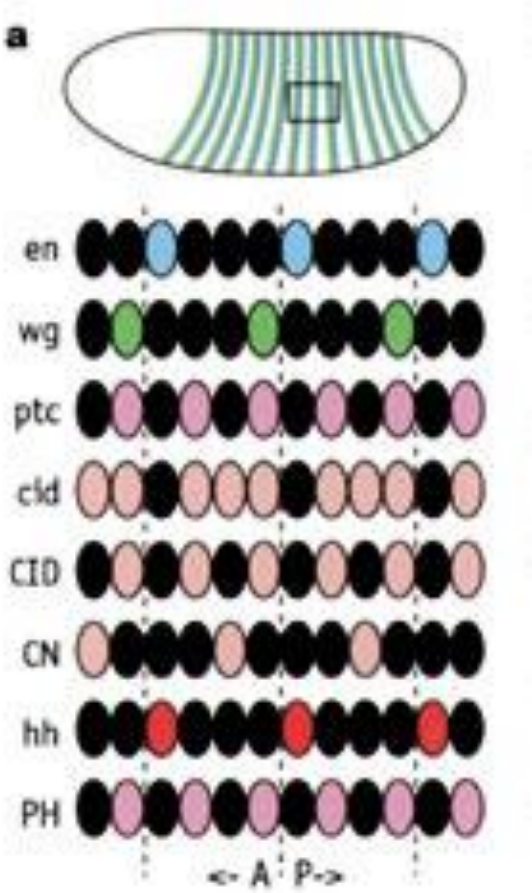
G. von Dassow et al., Nature 406, 188 (2000)

Gene expression patterns

The 2D pattern is reduced to 1D.
The gene expression is essentially binary (ON in some cells, OFF in others), and has a period of 4 cells.

Shown is the wild type steady state pattern of the segment polarity genes. Each row corresponds to an mRNA or protein.

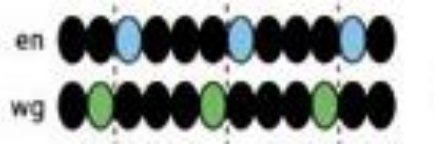
Black – no/low expression
Colored – moderate/high expression



G. von Dassow et al., Nature 406, 188 (2000)

Simulations

• Start from the wild type initial condition for *en* and *wg.* All other nodes are assumed to be not expressed.



• Generate a set of kinetic parameters from the biologically relevant range (48 unknown parameters)

• Run the simulation until steady state is reached.

• Use threshold (>6% of maximal concentration) to decide whether node is ON or OFF.

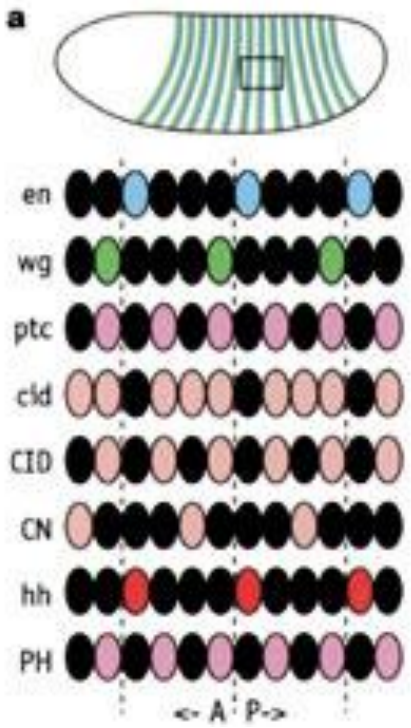
• Compare with wild type pattern, if the same accept as a solution

G. von Dassow et al., Nature 406, 188 (2000)

Gene expression patterns in the model

The 2D pattern is reduced to 1D, assume cells are hexagonal

wild type



model solution

Q: Can you spot differences between the real and model pattern?

Robustness to parameter changes

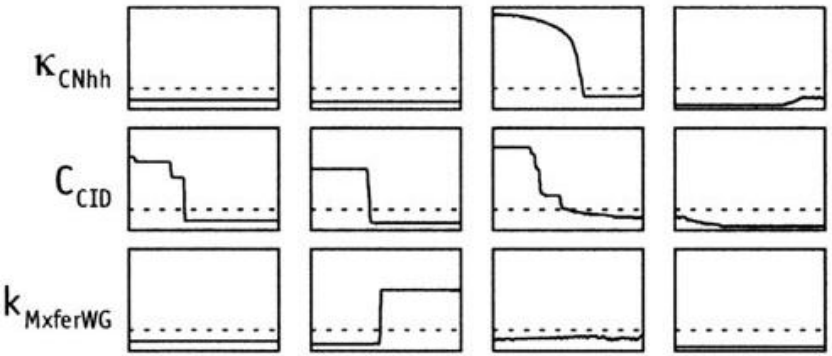
The von Dassow model has 13 equations and 48 unknown parameters.

$$\frac{d[hh]_i}{dt} = T_{\max}\rho_{hh}\left[\left[\frac{[EN]_i^{\nu_{ENhh}}}{K_{ENhh}^{\nu_{ENhh}} + [EN]_i^{\nu_{ENhh}}}\right]\right] - \frac{[hh]_i}{H_{hh}}$$

$$\frac{d[PH]_{i,j}}{dt} = k_{PTCHH}[HH]_{n,j+3}[PTC]_{i,j} - \frac{[PH]_{i,j}}{H_{PH}}$$

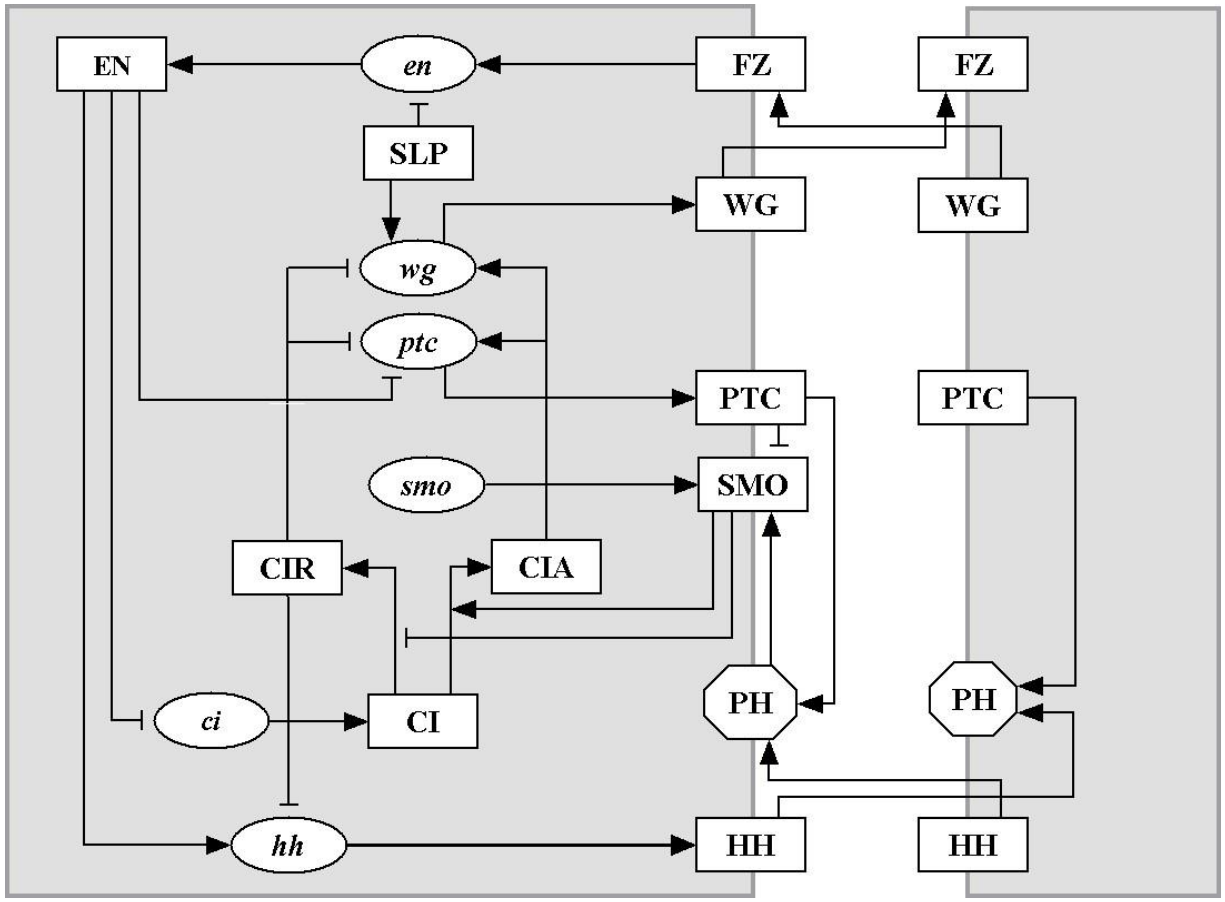
Systematic search shows that 1 in every 46 parameter combinations lead to wild type final patterns. The others are not good.

The parameter combinations leading to wild type steady states are distributed homogeneously in the biologically relevant parameter space.



It is not the fine-tuning of the kinetic rates but the overall network topology what matters.

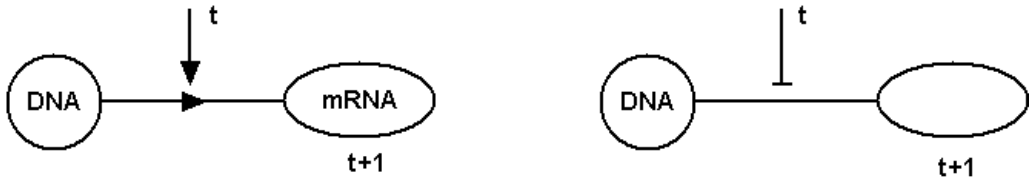
Segment polarity network in the Boolean model



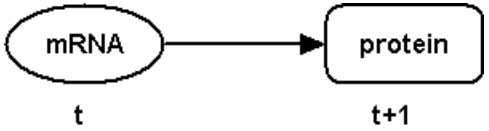
R. Albert, H. G. Othmer, Journ. Theor. Biol. 223, 1 (2003)

Synchronous Boolean model

- Transcripts and proteins are either ON (1) or OFF(0).
- The expression of a node at timestep *t* is given by a logical rule of the expression of its effectors at time *t-1*.
- Transcription depends on transcription factors; repressors are dominant.



- Translation depends on the presence of the transcript.



- Transcripts and most proteins decay if not produced
- Transcription, translation, mRNA/protein decay on the same timescale, protein binding much faster.

Rules for transcription and translation

$$en_i^{t+1} = (WG_{i-1}^t \text{ or } WG_{i+1}^t) \text{ and not } SLP_i^t$$

$$hh_i^{t+1} = EN_i^t \text{ and not } CIR_i^t$$

$$ptc_i^{t+1} = CIA_i^t \text{ and not } EN_i^t \text{ and not } CIR_i^t$$

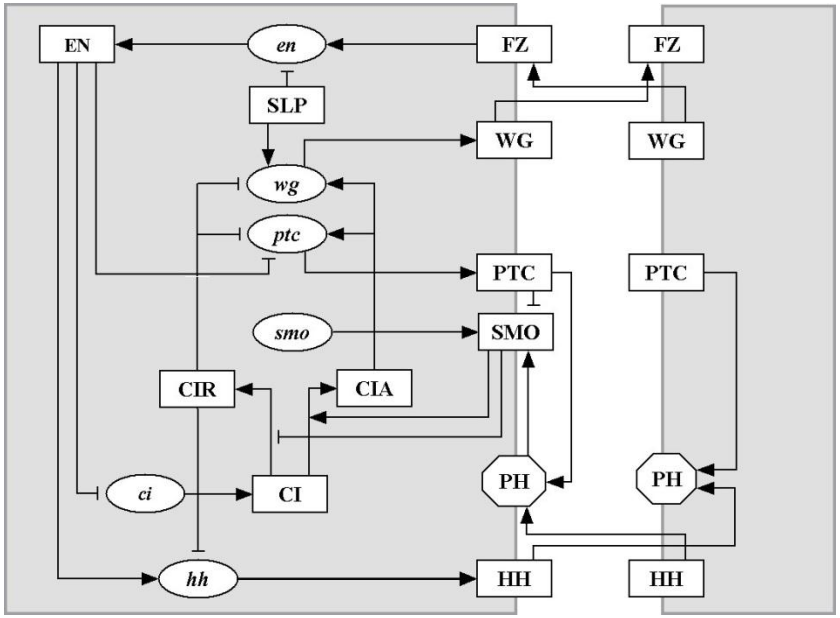
$$ci_i^{t+1} = \text{not } EN_i^t$$

$$EN_i^{t+1} = en_i^t$$

$$WG_i^{t+1} = wg_i^t$$

$$CI_i^{t+1} = ci_i^t$$

$$HH_i^{t+1} = hh_i^t$$



wg, PTC and SLP are more stable than other proteins

$$wg_i^{t+1} = (CIA_i^t \text{ and } SLP_i^t \text{ and not } CIR_i^t) \text{ or}$$
$$[wg_i^t \text{ and } (CIA_i^t \text{ or } SLP_i^t) \text{ and not } CIR_i^t]$$

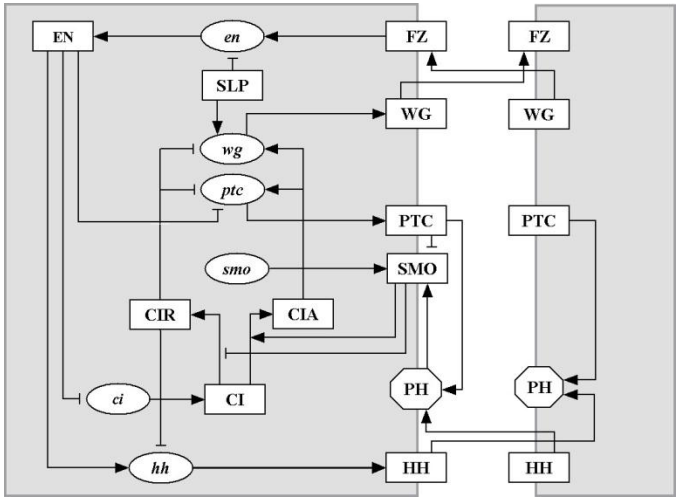
Either of the activators can counter mRNA decay.

$$PTC_i^{t+1} = ptc_i^t \text{ or } (PTC_i^t \text{ and not } HH_{i-1}^t \text{ and not } HH_{i+1}^t)$$

Free PTC does not decay .

$$SLP_i^{t+1} = SLP_i^t$$

SLP is a source in the segment polarity network.

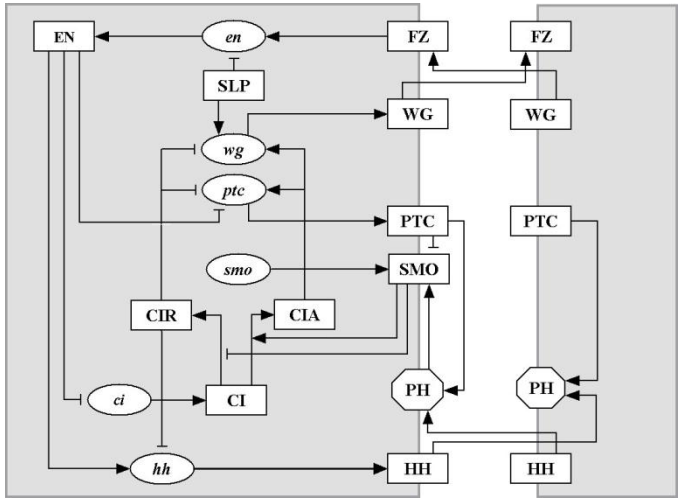


Rules for post-translational processes

$$PH_i^t = PTC_i^t \text{ and } (HH_{i-1}^t \text{ or } HH_{i+1}^t)$$

$$SMO_i^t = \text{not } PTC_i^t \text{ or } HH_{i-1}^t \text{ or } HH_{i+1}^t$$

instantaneous

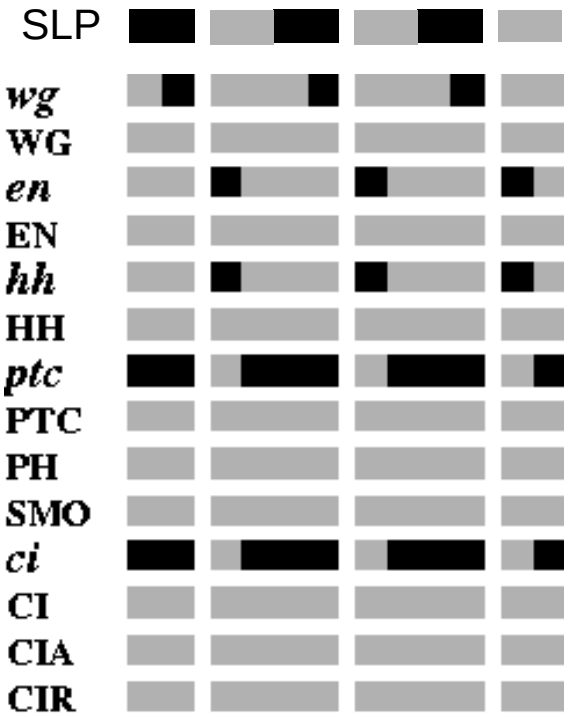


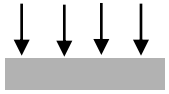
$$SMO_i^{t+\varepsilon}$$

$$CIA_i^{t+1} = CI_i^t \text{ and } (\overbrace{SMO_i^t \text{ or } hh_{i-1}^t \text{or } hh_{i+1}^t}^{SMO_i^{t+\varepsilon}})$$

$$CIR_i^{t+1} = CI_i^t \text{ and not } SMO_i^t \text{ and not } hh_{i-1}^t \text{and not } hh_{i+1}^t$$

Start the model from an initial state giving the prepattern of all nodes





Parasegment with four cells

Two cells of a previous/ next parasegment

cell with node ON

cell with node OFF

The patterns are assumed to be identical in each parasegment.

Wild type initial state: *wg* in the last cell of the parasegment, *en/hh* in the first cell of the parasegment, *ptc* and *ci* complementary to *en*, no proteins.

The model reproduces the wild type steady state

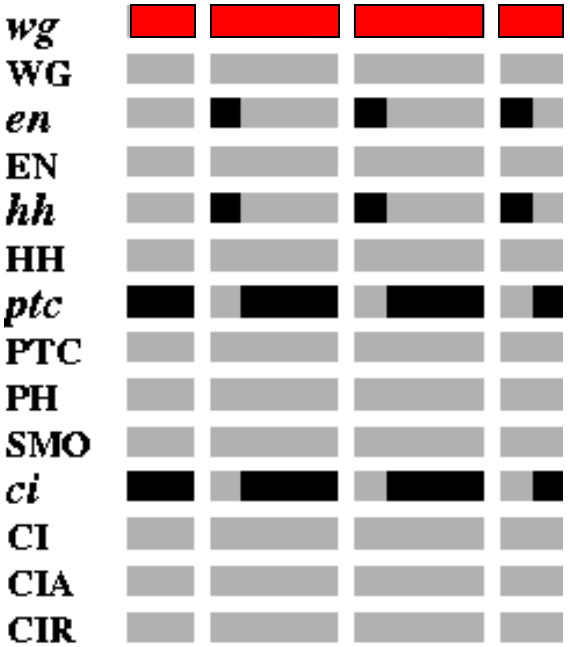
initial state

steady state

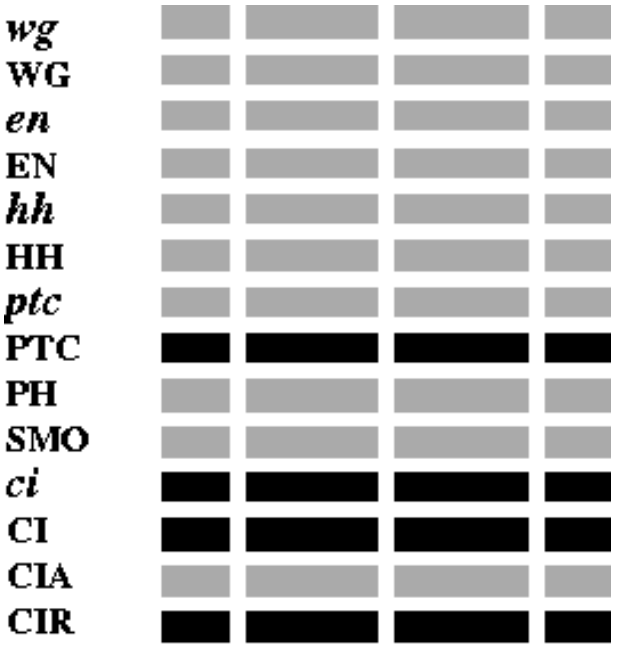
Agrees with the wild type pattern of the segment polarity genes.

The net effect of the interactions is enough to capture the functioning of the network!

wg, en or *hh* knockout mutations are lethal



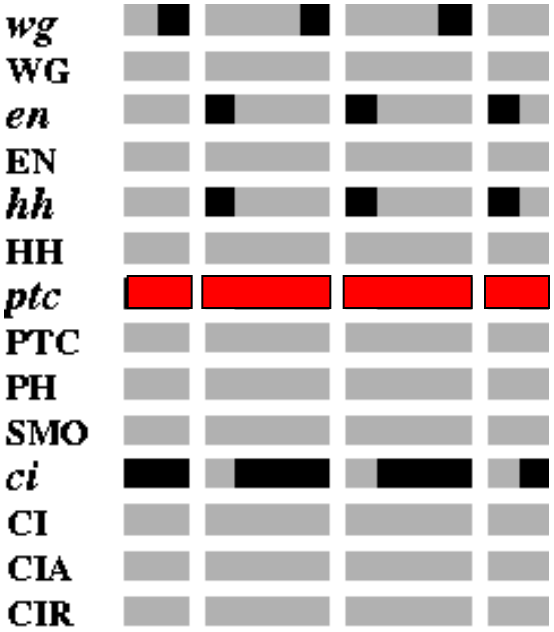
initial state



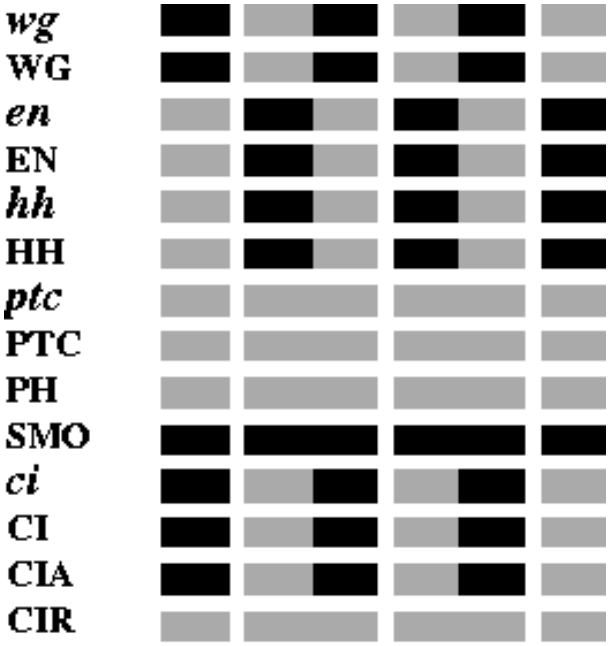
final state

Modeling knockout mutation: keep the node's state OFF (0).
Final state: no *wg*, *en* and *hh*, no segmentation, regardless of initial state.

ptc knockout broadens the stripes



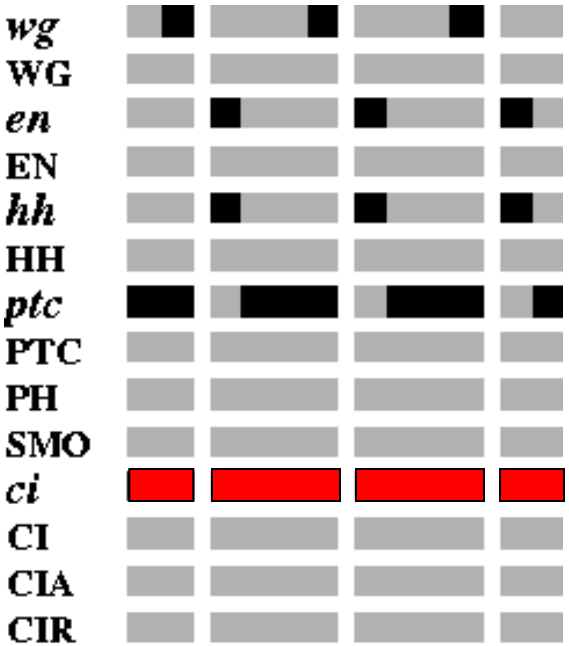
initial state



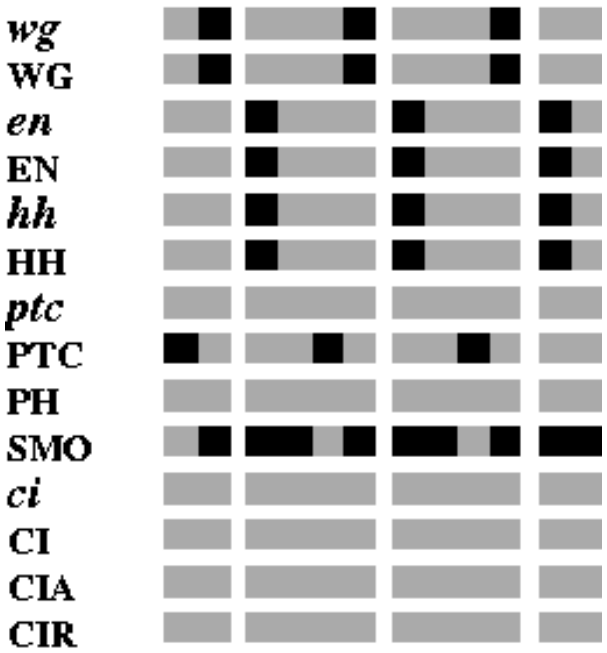
final state

The *wg*, *en* and *hh* stripes broaden, regardless of initial state.

ci knockout can preserve the prepattern



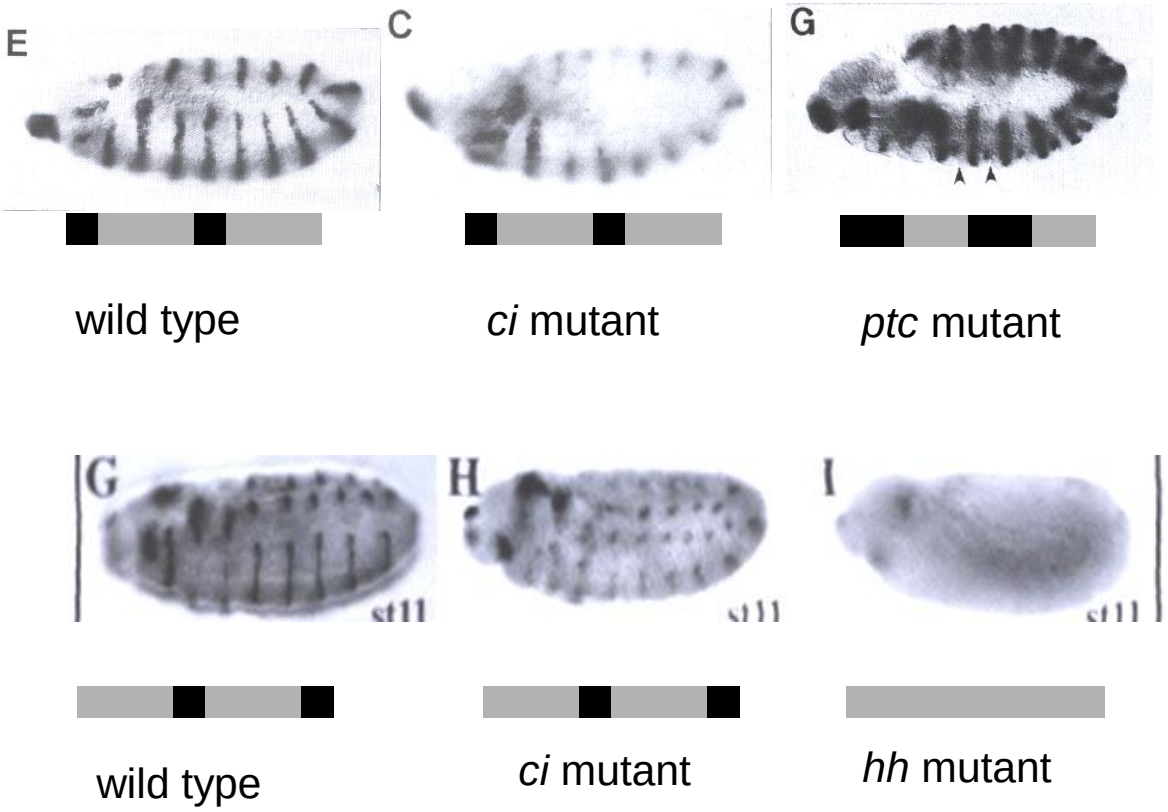
initial state



final state

The effect of *ci* mutation depends on the initial state.
For wild type prepattern, the *wg, en, hh* stripes remain.

Model correctly reproduces experimental results on knock-out mutants

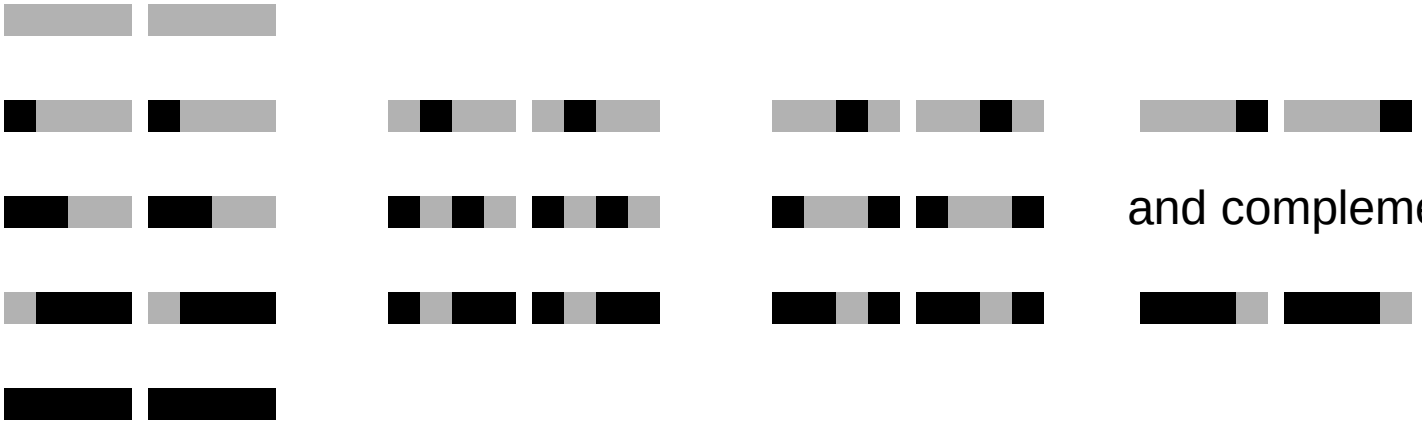


Tabata, Eaton, Kornberg, Genes & Development 6, 2635 (1992)

Gallet et al., Development 127, 5509 (2000)

Sensitivity to initial states

Possible number of prepatterns for a single node: $\mathbf{16}$

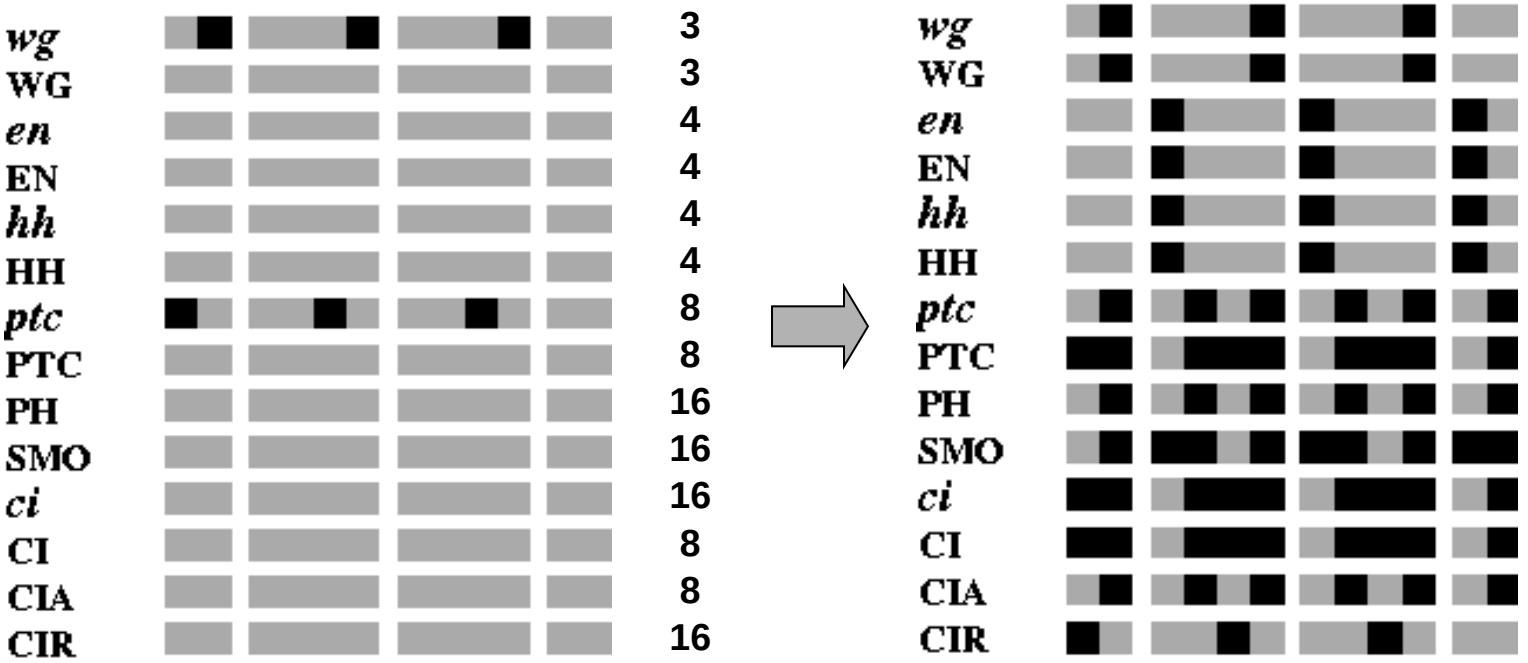


and complements

Total number of network-wide prepatterns: $N_i = 16^{15}$

All initial states lead to fixed points (steady states)
within 10-15 steps! - robustness

How many initial states lead to the wild type steady state?



minimal prepattern

steady state

Total number of wild-type inducing prepatterns: $6\times10^{11} = 8\times10^{-6}\, N_i$

The fixed points can be determined analytically

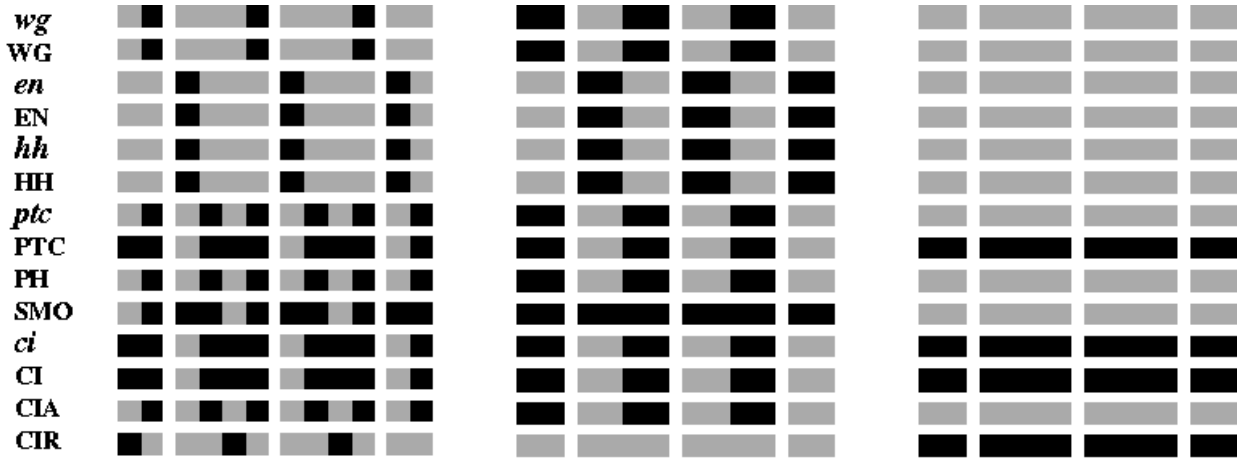
- In the stable state $x_i^{t+1} = x_i^t$
- Use the fact that $SLP_1 = SLP_2 = 0$ and $SLP_3 = SLP_4 = 1$
- The set of equations reduces to:

$$\begin{cases} wg_1 = wg_1 \text{ and not } wg_2 \text{ and not } wg_4 \\ wg_2 = wg_2 \text{ and not } wg_1 \text{ and not } wg_3 \\ wg_3 = wg_1 \text{ or } wg_3 \\ wg_4 = wg_2 \text{ or } wg_4 \end{cases}$$

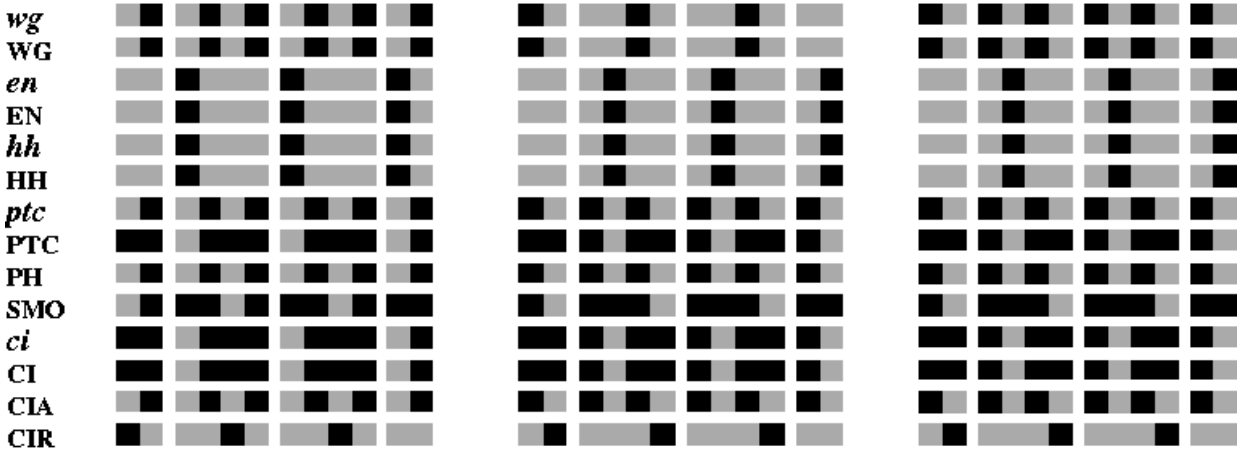
$$\begin{cases} PTC_1 = (\text{not } wg_2 \text{ and not } wg_4) \text{ or } (PTC_1 \text{ and not } wg_1 \text{ and not } wg_3) \\ PTC_2 = (\text{not } wg_1 \text{ and not } wg_3) \text{ or } (PTC_2 \text{ and not } wg_2 \text{ and not } wg_4) \\ PTC_3 = PTC_4 = 1 \end{cases}$$

Ex. Find the solutions of the first four (wg) equations.

Possible steady state patterns



wild type broad lethal



double *wg* displaced displaced, 2 *wg*

Asynchronous algorithms

1. different update time for each node

$$T_i^k = k\gamma_i$$

2. select a random update order in each timestep

3. assume that posttranslational processes are always faster than pre-translational ones

$$\gamma_{prot} < \gamma_{mRNA}$$

or, proteins always updated before mRNAs. In this case the rules of transcription and translation simplify, mixing is only possible for post-translational rules

e.g. $hh_i^t = EN_i^t \quad \text{and not } CIR_i^t \qquad EN_i^{t+1} = en_i^t$

Explore all possible durations

Start from the wild type initial state.

Select the update times or update order uniformly randomly
- Update times stay the same
- Randomize the update order in each step.

The steady states of the model are the same as the synchronous model's, but now oscillations are also possible.

The system is not deterministic anymore: the same initial condition can lead to different steady states, depending on the order of update.

Determine the incidence (probability) of the different steady states.

Asynchronous update causes that the WT initial state leads to a variety of steady states

Steady state	Incid.
wild type	60%
broad	24%
lethal	15%
displaced	1%

random order

Steady state	Incid.
wild type	87.5%
broad	12.5%
lethal	0%
displaced	0%

Two-timescale

M. Chaves, R. Albert, E. Sontag, Journ. Theor. Bio. 235, 431 (2005).

Continuous-Boolean hybrid model

- Each node is characterized by both a continuous and a Boolean variable.

$$\frac{d\hat{X}_i}{dt} = \alpha_i \left(B(X_{1,} X_2 ..) - \hat{X}_i \right)$$

- X_i is defined by the threshold rule

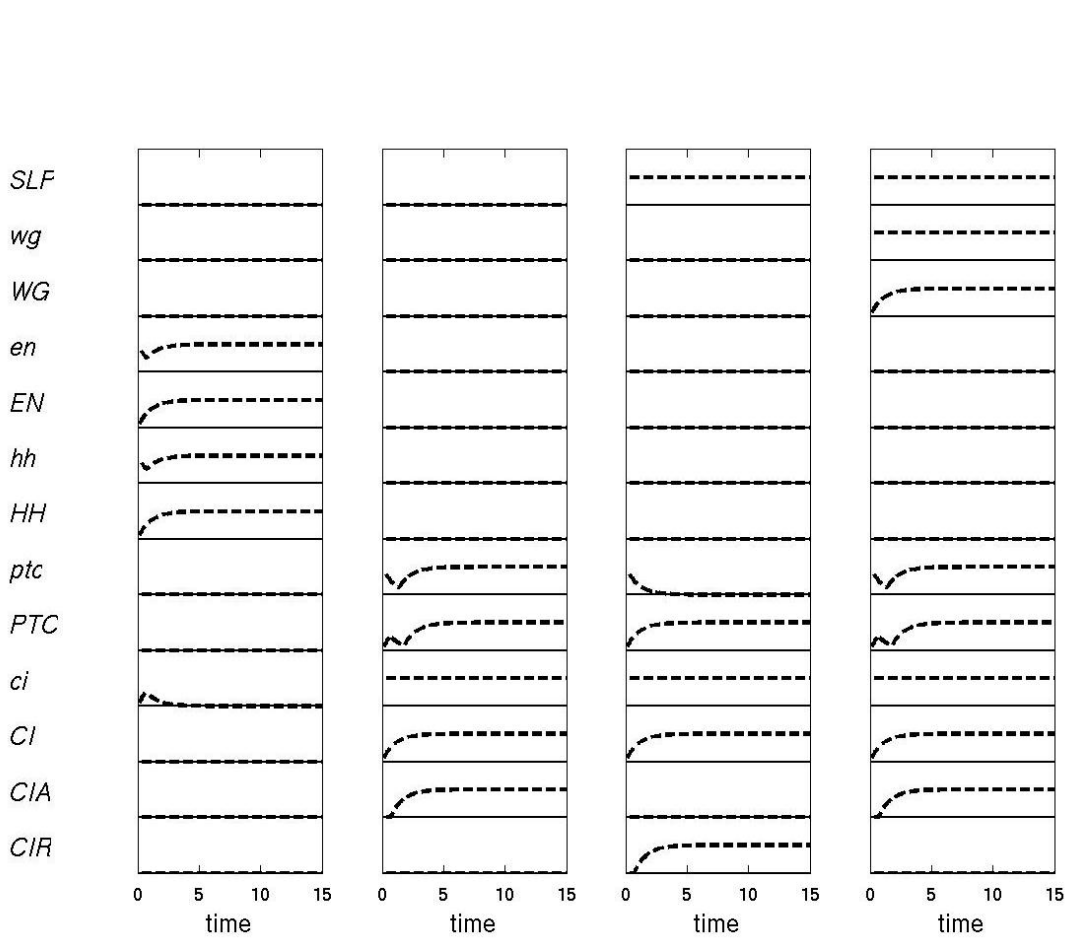
$$X_i = \begin{cases} 0, & \text{if } \hat{X}_i < \theta_i \\ 1, & \text{if } \hat{X}_i \geq \theta_i \end{cases}$$

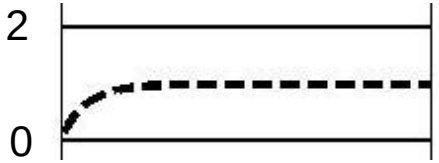
- Compared to $\frac{dX}{dt} = T_{\max} \rho_X \left(\frac{Y^v}{K_Y^v + Y^v} \right) - \frac{X}{H_X}$, this assumes

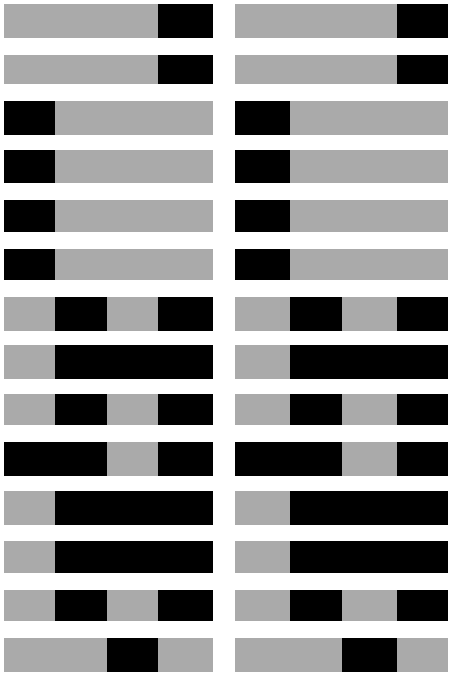
maximal synthesis rate = decay rate= α_i

M. Chaves, E. Sontag, R. Albert, IEE Proc. Systems Biology (2006)

Wild type behavior is reproducible







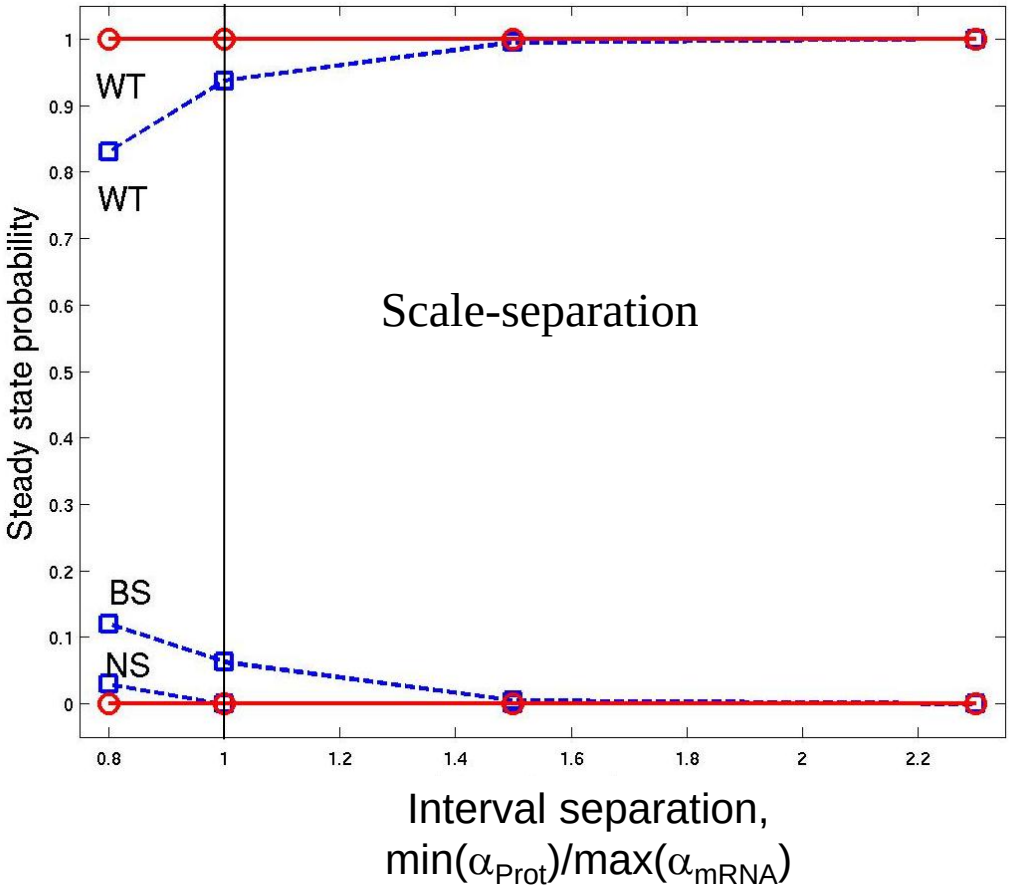
Hybrid model more robust than asynchronous Boolean model

- Scale-separation: choose the scale factors from A_{mRNA} and A_{prot}
- Faster protein synthesis/decay
- Start from WT initial condition, calculate incidence of steady states

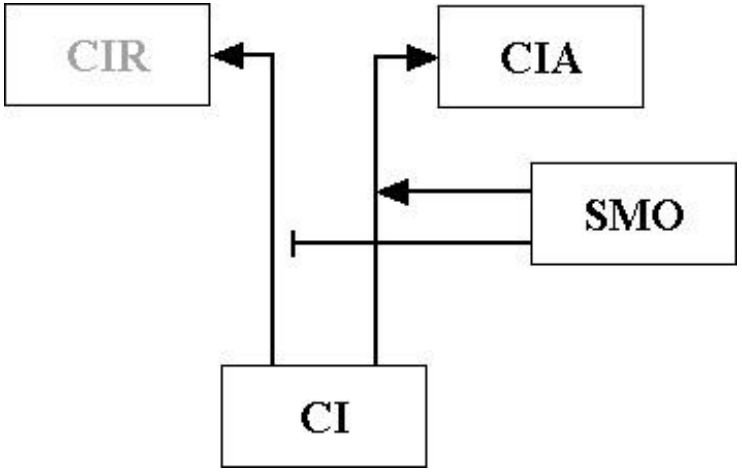
hybrid model

asynchronous

If protein scale factors disjoint from mRNA scale factors, the only possible steady state of the hybrid model is the WT.



Can we find the cause of the divergence?



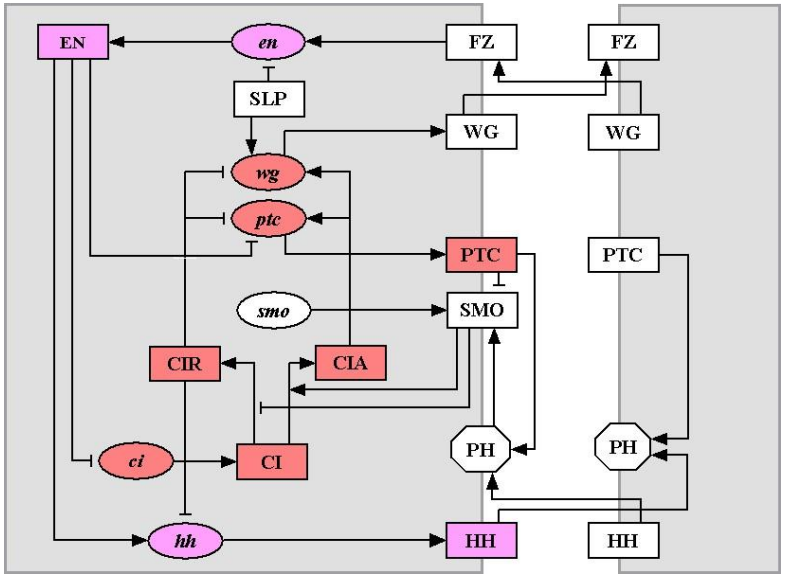
The two CI transcription factors have opposite regulatory roles

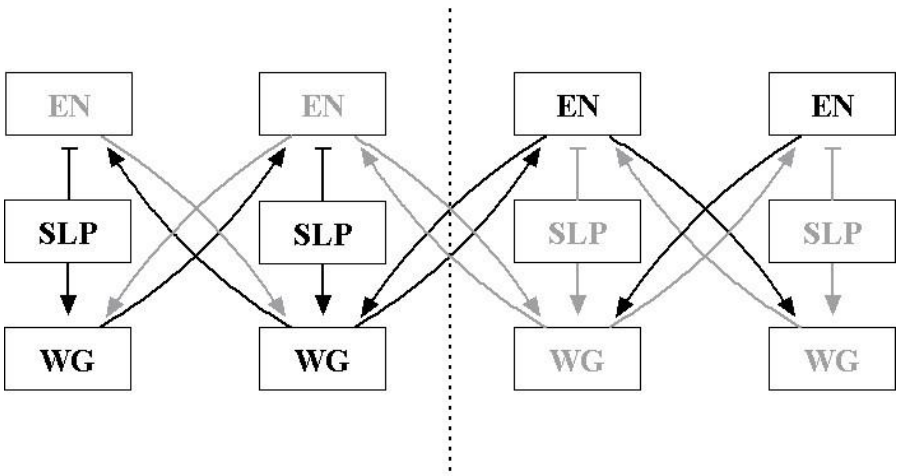
The post-translational modification of CI is regulated in a binary fashion.

- The expression of CIA and CIR needs to be complementary in all CI-expressing cells
- If a perturbation of cellular process durations leads to an imbalance between CIA and CIR, the wild type steady state becomes unreachable.
- Only CIA - broad stripes; only CIR - no segmentation
- The condition of CIA/CIR complementarity in the two-timescale algorithm is that PTC be initiated before SMO – true biologically

Interplay between topology and function

- The network contains two activating clusters that inhibit each other in each cell, *en, hh* and *ci, wg, ptc*
- At the same time *en* and *wg* enforce each other in neighboring cells through the secreted proteins HH and WG





- SLP is a regulatory source that maintains asymmetry and limits *en* and *wg* to different halves of the parasegment.